

APPLICATION
FOR
UNITED STATES LETTERS PATENT

TITLE: ROUTING DATA

APPLICANTS: FARUK S. HADZIOMEROVIC

Express Mail No.: EL594060565US
Date: October 16, 2000

Prepared by: Trop, Pruner & Hu, P.C.
8554 Katy Freeway, Ste. 100, Houston, TX 77024
713/468-8880 [Office], 713/468-8883 [Fax]

ROUTING DATATECHNICAL FIELD

The invention relates generally to networks, and, more particularly, to routing signals efficiently over a network.

5

BACKGROUND

Communication networks are widely used to link various nodes, such as personal computers, servers, gateways, and the like. Communication networks may include private networks, for example, local area networks (LANs) and wide area networks (WANs), or public networks as in the Internet. Popular forms of communications across such networks include electronic mail, file transfer, web browsing, and other exchanges of data.

With the increased capacity and reliability of communications networks, transmission of real-time information, such as voice and streaming video, over networks has become possible. The ability to transfer a wide variety of information has made communication networks a popular choice of communication amongst users, as evidenced by the emergence of the Internet, for example. The popularity of networks as an attractive mode of communication has resulted in increased traffic, thereby making today's networks more prone to congestion, particularly during peak hours.

To avoid network congestion, developers routinely employ routing algorithms to determine efficient paths from source to destination nodes. Generally, routing algorithms perform two main functions -- selecting routes between the source and destination nodes and delivering messages to the

intended destination after the routes have been determined. The latter function of delivering messages to the intended destination involves use of data structures known as routing tables to deliver messages. The former function of selecting routes typically involves a collection of routing algorithms executing at different nodes that exchange services or perhaps information.

It is desirable that routing algorithms correctly and expeditiously determine efficient routes to transmit data. Inefficient routing may delay the delivery of data, or even cause data to be lost, which may lead to network congestion.

A need thus exists for an improved method and apparatus of routing data within a communications system.

15 SUMMARY

In general, according to one embodiment, a method comprises storing a first data structure containing costs associated with transmitting data between routers in a network. The method comprises combining the first data structure with itself to determine a cost of transmitting the data. The method further comprises transmitting the data along a route based on the calculated cost.

Some embodiments may have one or more of the following advantages. A reliable and efficient method and apparatus are provided for routing data from a source to a destination node. In one embodiment, a data packet may be transmitted to its intended destination in a "cost" efficient manner, where the cost may be one or more of the factors including physical distance, delay, expense, reliability, bandwidth, load, and security between nodes.

Other features and advantages will become apparent from the following description, from the drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

5

Figure 1 is a block diagram of an embodiment of a communications system in accordance with the present invention;

10 Figure 2 is a block diagram of a router that may be employed in the communications system of Figure 1;

15 Figure 3 illustrates an exemplary data packet 250 that may be communicated in the communications system of Figure 1;

20 Figure 4 is a flow diagram of a method in accordance with one embodiment of the present invention that may be implemented in the communications system of Figure 1;

25 Figure 5 is a flow diagram of a method in accordance with one embodiment of the present invention that may be implemented in the communications system of Figure 1; and

Figure 6 is a flow diagram of a method in accordance with one embodiment of the present invention that may be implemented in the communications system of Figure 1.

DETAILED DESCRIPTION

30 In the following description, numerous details are set forth to provide an understanding of the present invention. However, it will be understood by those

skilled in the art that the present invention may be practiced without these details and that numerous variations or modifications from the described embodiments may be possible.

5 Referring now to Figure 1, an exemplary communications system 1 comprising a network 3 in accordance with one embodiment of the present invention is illustrated. In one embodiment, the network 3 may be a packet-based network, such as an Asynchronous Transfer Mode (ATM) network or an Internet Protocol (IP) network. As utilized herein, the term "network" may be 10 one or more communications networks, channels, links, or paths and systems or devices (such as routers) used to route data over such networks, channels, links, or paths.

15 The ATM standard is established by the International Telegraph and Telephone Consultative Committee Communications (CCITT) Standards Organization. One version (Version 4) of IP is described in Request for Comments (RFC) 791, entitled "Internet Protocol," dated September 1981. Another version of IP is IPv6, as described in RFC 2460, entitled "Internet 20 Protocol, Version 6 (IPv6) Specification," dated December 1998. In an ATM network, packets or cells are used to communicate information.

25 Although not so limited, in the illustrated embodiment, the network is an IP network. The IP network is a connectionless packet-switched network in which routing of packets is based on a source and destination address carried in each packet. The packets can travel independently over any path (and possibly over different paths) to a destination point identified by the destination address carried in the packets. Thus, in an IP network, the packets may even arrive out of order.

The network 3 comprises two hosts 4-5 and six nodes 6-11. In one embodiment, the hosts 4-5 may be computers that are part of local area networks. The six nodes 6-11 are linked to each other directly, or through one or more nodes 6-11. For ease of illustration, only six nodes 6-11 are illustrated, 5 although the network 3 may include fewer or more than six nodes. Similarly, the network 3 may include additional or fewer than two hosts 4-5. In one embodiment of the invention, each node 6-11 includes a router 26-31, which may be implemented in a computer, server, gateway, internet device, and the like. Internet devices may include televisions, appliances, or any other device 10 capable of interfacing with the network 3.

Each router 26-31 determines the next network node to which a data packet is forwarded before the data packet ultimately arrives at its final destination. As described more fully below, in accordance with one embodiment 15 of the instant invention, one or more of the routers 26-31 in the network 3 determine an efficient route to deliver data packets to the final destination. A Router 26-31 may be located at any junction of networks or in gateways, including at an Internet point of presence (POP). A POP is the access point to 20 the Internet, and has a unique Internet address. Independent service providers or online service providers typically have a point of presence on the Internet.

Routers 26-31 typically create or maintain a table of the available routes and the route conditions and use this information to determine the best route for 25 a given packet. The packet may travel through a number of network nodes, and hence through several routers, before arriving at its final destination.

As shown, the first node 6 is coupled to the second node 7 over a first link 64, to the third node 8 over a second link 66, and to the fourth node 9 over a third link 68. The second node 7 is coupled to the third node 8 over a fourth link 30 70 and the fourth node 9 over a fifth link 72. The third node 8 is coupled to the

fifth and sixth nodes 10, 11 over the sixth and seventh links 74, 76, respectively. The fourth node 9 is coupled to the second node 2 over the fifth link 72, to the third node 8 over an eighth link 76, and to the fifth node 10 over a ninth link 78. The fifth node 10 is coupled to the sixth node 11 over an eleventh link 80. For 5 illustrative purposes, the fourth link 72 is depicted as a bi-directional link, while the other links 64, 66, 68, 70, 74, 76, 78, 80 are unidirectional; however, it should be appreciated that the network 3 may include any combination of unidirectional or bi-directional links.

10 Each link may have a "cost" associated with it. The "cost," as used herein, may, in one embodiment, represent one or more of the following: physical distance, delay, expense, reliability, bandwidth, load, or security between nodes. Physical distance is typically the sum of the distance associated with each link traversed, or, alternatively, may be defined in terms of "hop 15 count," which refers to the number of passes through internetworking products, such as routers, that a packet takes en route from a source to a destination node. Reliability typically refers to the dependability (usually described in terms of the bit-error rate) of each network link. Routing delay, which may depend on the bandwidth of intermediate network links, the port queues at each router along the way, network congestion on all intermediate network links, and the 20 like, usually refers to the length of time required to move a packet from a source to destination node. Bandwidth typically refers to the available traffic capacity of a link, while load refers to the degree to which a network resource, such as a router, is busy. Load can be calculated in a variety of ways, including central 25 processing unit (CPU) utilization and packets processed per second. Communication expense refers commonly to the expenditure associated with transmitting data packets from a source to destination node. Security typically refers to the level of security associated with each link or path. In alternative embodiments, the "cost" may be any other standard of measurement upon

which it is desirable to determine routing. The cost may be different for different directions on the same link.

For illustrative purposes, exemplary costs are shown between the nodes 5 6-11 of the network 3. Although not so limited, the cost associated with each link in the illustrated embodiment represents the physical distance between any two particular nodes. For instance, the cost between the first node 6 and second, third, and fourth nodes 7, 8, 9, is two, five, and one, respectively. The cost between second node 7 and third and fourth nodes 8, 9 is three and two, 10 respectively. The cost between the third node 8 and the fourth, fifth, and sixth nodes 9, 10, 11, is three, one, and five, respectively. The cost between the fourth and fifth nodes 9, 10 is one, and the two between the fifth and sixth nodes 10, 11.

15 In accordance with one embodiment of the instant invention, one or more of the routers 26-31 determine a "cost" efficient route for data packets between a source and destination node. Figure 2 depicts one embodiment of a router 200 in accordance with the present invention. While the router 200 of Figure 2 may be implemented at any or all of the nodes 6-11 in the network 3 of Figure 1, for ease of illustration it is herein assumed that the router 200 represents the first router 26 at the first node 6 of the network 3.

20 The router 200 comprises first, second, and third interfaces 205, 210, 215, that are respectively coupled to the routers 27, 28, 29 at the second, third, and fourth nodes 7, 8, 9 (see Figure 1), respectively. A fourth interface 216 is provided for interfacing with the first host 4. The router 200 comprises an input queue 217 for receiving data packets from the first, second, third and fourth interfaces 205, 210, 215, 216. An exemplary structure of an IP data packet 250 is illustrated in Figure 3. The data packet 250 comprises an IP version field 252 that indicates the current IP version. A header length field 254 is provided for

identifying the length of the header of the IP data packet 250. The data packet 250 comprises a type of service field 256 that allows the data packet 250 to be delivered based on a variety of service factors, such as delay, throughput, reliability, expense, and the like. Thus, it is possible to transmit the data packet 5 250 through a route that presumably has the least amount of delay, highest throughput, least chance of getting lost, or least cost.

The data packet 250 comprises a total IP length field 258 for indicating the total length of the data packet 250, including the IP header and the 10 accompanying data. The maximum length for the IP data packet 250, for example, is 65,535 bytes (for IPv4). The data packet 250 may also include other fields, such as the packet ID number, fragment, and time-to-live, and IP header checksum fields 260, 262, 264, 266. A protocol carried field 268 identifies a higher layer protocol, such as Transmission Control Protocol (TCP), User 15 Datagram Protocol (UDP), Internet Control Message Protocol (ICMP), and the like.

The data packet 250 comprises source and target address fields 270, 272. In Version 4 of the Internet Protocol, the source and target address fields 270, 20 242 of the data packet 250 are typically 4 bytes long. An options field 274 is provided in the data packet 250 for enabling measurement of time delay, routing messages, or specifying routing instructions. The options field 274, for example, may be utilized to specify data packet 250 route from source to destination by providing the intermediate routers that the data packet 250 should traverse. In 25 Version 4 of the IP, for example, up to nine (9) intermediate routers may be specified between the source and destination. The data packet 250 also comprises a data field 276 that contains the data being transmitted.

Referring again to Figure 2, the incoming data packets 250 are provided 30 from the input queue 217 to an IP options block 295, which processes the

options field 274 of each received data packet 250 to determine if source routing is specified. In alternative embodiments, the IP options block 295 may perform additional functions as well. If source routing is specified in the IP options field 274, the data packet 250 is provided to the IP output block 300, which forwards 5 the data packet 250 to the next router specified in the options field 274 through the interfaces 205, 210, 215.

If the IP options block 295 determines source routing is not specified, then, at block 315, the router 200 determines if the data packet 250 is intended 10 to be forwarded, broadcast, or delivered to the first host 4 coupled to the router 200. If the data packet 250 is to be forwarded or broadcast to other routers 27-31 on the network 3, then the data packet 250 is provided to the IP output block 300, which transmits the data packet 250 to the appropriate router or routers 15 27-31. If the data packet 250 is not intended to be forwarded or broadcast to other routers 27-31, it may be provided to a TCP or UDP layer for delivery to the first host 4 through the interface 216, or, alternatively, it may be delivered to an Internet Control Message Protocol (ICMP) block 325. The ICMP block 325 typically handles error and control messages for IP.

The router 200 includes a routing table 330 that in one embodiment 20 contains information to aid the process of path determination. Routing information stored in the routing table 330 may vary, depending on the routing algorithm employed. The routing table 330 may include destination (next hop) associations that indicate to the router 200 that a particular destination may be reached efficiently by transmitting the data packet 250 to a particular router 25 representing the "next hop" on the way to the final destination. In one embodiment, the routing table 330 may contain information about the "costs", in terms of delay, distance, and the like, that are associated with transmitting data packets 250 between routers.

Routers 26-31 (see Figure 1) routinely communicate with one another and exchange messages to build a substantially detailed picture of network topology. The routing table 330 of the router 200 may be updated by a route command 340, routing daemon 335, or an ICMP redirect 325. The route command 340 to update the routing table 330 may be issued by an operator sitting at the host 4, for example. The routing daemon 335 is a software application that is capable of executing Routing Information Protocol (RIP), which is commonly used to periodically exchange routing information from other computers on the network 3 (e.g., about every 30 seconds).

10

In accordance with one embodiment of the present invention, a route calculation block 345 of the router 200 uses the information from the routing table 330 to calculate a "cost" efficient route for the packet data 250. One embodiment of a routing algorithm employed by the route calculation block 345 is described in more detail below. The route calculation block 345 determines an efficient route for delivering the packet data 250 and provides the routing information to the IP output block 300, which then forwards the packet data 250 to the next router based on the route determined by the route calculation block 345.

20

Selected portions of the router 200, such as the routing calculation block 345, may be implemented in hardware, software, or a combination thereof. In one embodiment, the router 200 may include software routines, modules, or drivers that may be executable on the control unit 245. Data instructions associated with the software routines may be stored in the storage unit 360.

Referring now to Figure 4, one method in accordance with the present invention of delivering data packets 250 efficiently over the network 3 based on one or more of the "costs" is illustrated. For illustrative purposes, the method of Figure 4 is described in the context of transmission of the data packet 250 from

the first host 4, which is coupled to the first router 26, to the second host 5, which is coupled to the sixth router 31. The method of Figure 4 begins at block 450, where the first router 6 receives the data packet 250 from the first host 4 for delivery to the second host 5. Accordingly, the source address field 270 5 includes the IP address of the first host 4, and the target address field 272 includes the IP address of the second host 5. At block 460, the first router 26, based on the type of service field 256 of the data packet 250, determines the type of service that has been designated by the first host 4. That is, the first host 4 may, in one embodiment, designate that the data packet 250 be delivered 10 to the second host 5 using the shortest route. In alternative embodiments, the first host 4 may designate that the packet data 250 be delivered using the least expensive route or the most reliable route, for example.

At block 470, the first router 26 calculates an efficient route based on the type of service designated by the first host 4. Figure 5 illustrates one embodiment of the block 470. The method of Figure 5 begins at block 510, where the route calculation block 345 of the first router 26 generates a 15 "connection" data structure of the network 3. Although a variety of data structures may be utilized, in the illustrated embodiment the data structure utilized herein is a matrix. A connection matrix is provided that is representative 20 of the network topology, where the connection matrix illustrates the adjacent nodes to each node. An exemplary connection matrix for the network 3 of Figure 1 is as follows:

$$\text{Connection Matrix } C = \begin{vmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

Since the network 3 of Figure 1 comprises six nodes 6-11 (see Figure 1), the above connection matrix has six rows and six columns. Each row of the connection matrix illustrates the adjacent nodes of each node. Numerals "1" and "0" denote whether nodes are adjacent to each other, in particular, numeral "1" denotes that two particular nodes are adjacent, while numeral "0" indicates no adjacency. For example, the first row, which comprises numerals "1 1 1 1 0 0," indicates that the first node 6 is adjacent to the second, third, and fourth nodes 7, 8, 9, but not to the fifth and sixth nodes 10, 11. Likewise, the second row, which comprises sequence "0 1 1 1 1 0 0," indicates that the second node 7 is adjacent to the third and fourth nodes 8, 9, but not to the first, fifth and sixth nodes 6, 10, 11. Similarly, rows third, fourth, fifth, and sixth rows depict the adjacent nodes for respective third, fourth, fifth, and sixth nodes 8-11.

The route calculation block 345 (see Figure 2) may generate the data for the connection matrix in a variety of ways, including generating based on the information stored in the routing table 330. The routing table 330, in one embodiment, may include information regarding nodes that are adjacent to the first router 26. For example, the first router 26 may obtain routing table information about its adjacent nodes by submitting a layer 1 (Ethernet) broadcasting request to each of its interfaces 205, 210, 215. The routing table 330 may use routing protocols, such as Open Shortest Path First (OSPF) and Routing Information Protocol (RIP), to obtain data for the other rows. For example, the first router 26 may utilize "flooding" to acquire data for the other rows of the connection matrix. Flooding involves each router 26-31 periodically broadcasting its adjacent nodes to all its neighboring nodes. The neighboring node in turn forwards the received broadcast message, along with the information about its own neighboring nodes, to all of its neighbors. Accordingly, the routing table 330, in one embodiment, comprises information about the adjacent nodes of the first router 26, as well as about the adjacent nodes of other nodes in the network 3. The route calculation block 345, in one

embodiment, accesses the routing table 330 to generate the connection matrix for the network 3.

At block 520, the route calculation block 345 generates a "cost" matrix.

5 In the illustrated embodiment, the cost matrix is representative of the distance between nodes in the network 3. An exemplary cost (*e.g.*, distance) matrix for the network 3 of Figure 1 is depicted below:

$$\text{Distance Matrix } D = \begin{vmatrix} 1 & D^2 & D^5 & D & \infty & \infty \\ \infty & 1 & D^3 & D^2 & \infty & \infty \\ \infty & \infty & 1 & \infty & D & D^5 \\ \infty & D^2 & D^3 & 1 & D & \infty \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix}$$

10

As shown, although not so limited, in the illustrated embodiment, the cost factor is expressed as exponents for reasons explained in more detail later. In the illustrated embodiment of the cost matrix, the letter 'D' is a variable (without value) that stands for distance. The cost matrix has six rows and six columns since the network 3 of Figure 1 has six nodes 6-11. Each row of the cost matrix depicts the minimum distance between a particular node and the remaining nodes in one step (*i.e.*, hop). That is, for example, the first row depicts the minimum one-step distance between the first node 6 and the remaining nodes 7-11, while the second row depicts the minimum one-step distance between the second node 7 and the remaining nodes 1, 8-11, and so forth.

15

20

The first row, which comprises the sequence "1 D² D⁵ D ∞ ∞," indicates that the minimum distance (expressed as exponents) in one step between the first node 6 and second, third, fourth nodes 7, 8, 9, is respectively 2, 5, and 1.

25 The minimum one-step distance between the first node 6 and fifth and sixth

nodes 10, 11 is infinity, since no direct connection exists between the first node 6 and fifth and sixth nodes 10, 11. The first element of the first row represents the distance between the first node 6 to itself. In this case, the first element of the first row is one because expressing the distance (*i.e.*, zero) between the first node 6 and itself as an exponent yields one (*i.e.*, D^0 equals 1). Similarly, the fourth row the cost matrix, which comprises the sequence “ ∞ D^2 D^3 1 D ∞ ,” illustrates that the minimum one-step distance between the fourth node 9 and the second, third, and fifth nodes 7, 8, 10 is 2, 3, 1, respectively. The fourth row also illustrates that no one-step connection exists between the fourth node 9 and first and sixth nodes (as denoted by the symbol ∞) 6, 11, while the distance between the fourth node 9 and itself is zero (denoted as one since D^0 equals 1).

15 The route calculation block 345 may, in one embodiment, generate data for the cost matrix in a similar way as the data for the connection matrix is generated. That is, flooding may be utilized to gather data for the cost matrix between each of the neighboring nodes in the network 3. Any other available means for obtaining the data for the cost matrix may be employed without deviating from the spirit of the instant invention.

20 At block 530, the route calculation block calculates a next_step matrix. The next_step matrix in one embodiment is calculated by combining the cost matrix with itself. The cost matrix is combined with itself in one embodiment using the following mathematical function:

25
$$\text{next_step cost matrix } (D^2_{ij}) = \min_{k=1 \text{ to } 6} (D_{ik} * D_{kj}),$$

30 where i indicates the row and j indicates the column of the cost matrix (D), and k ranges from 1 to the number of nodes (*e.g.*, six in the illustrated embodiment) in the network. The cost matrix, as described above, illustrates the minimum one-step distance between node i (row) and node j (column). In accordance with

one embodiment of the present invention, combining the cost matrix with itself once using the above mathematical function yields a resultant matrix that illustrates the minimum distance between a particular node and the other nodes using two or fewer steps. The resultant next_step data structure in the illustrate embodiment is:

$$D^2_{ij} = \begin{vmatrix} 1 & D^2 & D^5 & D & \infty & \infty \\ \infty & 1 & D^3 & D^2 & \infty & \infty \\ \infty & \infty & 1 & \infty & D & D^5 \\ \infty & D^2 & D^3 & 1 & D & \infty \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} \begin{vmatrix} 1 & D^2 & D^5 & D & \infty & \infty \\ \infty & 1 & D^3 & D^2 & \infty & \infty \\ \infty & \infty & 1 & \infty & D & D^5 \\ \infty & D^2 & D^3 & 1 & D & \infty \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} = \begin{vmatrix} 1 & D^2 & D^4 & D & D^2 & D^{10} \\ \infty & 1 & D^3 & D^2 & D^3 & D^8 \\ \infty & \infty & 1 & \infty & D & D^3 \\ \infty & D^2 & D^3 & 1 & D & D^3 \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix}$$

The above mathematical function is similar to ordinary matrix multiplication except, instead of summation, the minimum function is employed. As an example, to obtain the first element (*i.e.*, first row, first column), the first row ("1 D² D⁵ D ∞ ∞") of the cost matrix is multiplied with the first column ("1 ∞ ∞ ∞ ∞ ∞") of the cost matrix, and instead of summing each element, the smallest (*i.e.*, minimum) element is selected as the first element. That is, the smallest element is selected from the following elements [*1*, (*i.e.*, 1x1), ∞ (*i.e.*, D² x ∞), ∞ (*i.e.*, D⁵ x ∞), ∞ (*i.e.*, D x ∞), ∞ (*i.e.*, ∞ x ∞), ∞ (*i.e.*, ∞ x ∞)], which in the instant example is one (1). As an added example, to obtain the element at row 2, column 4, the second row ("∞ 1 D³ D² ∞ ∞") of the cost matrix is multiplied with the fourth column ("D D² ∞ 1 ∞ ∞") of the cost matrix, and the smallest element from the following elements [∞ (*i.e.*, ∞ x D), D² (*i.e.*, 1 x D²), ∞ (*i.e.*, D³ x ∞), D² (*i.e.*, D² x 1), ∞ (*i.e.*, ∞ x ∞), ∞ (*i.e.*, ∞ x ∞)] is selected. In this case, the smallest element is D². Other elements of the D²_{ij} cost matrix are calculated similarly.

Since the elements of the matrix are either exponents of D or infinity, the multiplication reduces to exponent summation, which generally is a relatively quick and simple operation. Furthermore, since the minimum distance between any two nodes can only be one, the mathematical operation over k iterations can

stop as soon as a resulting element of D having a value of one is found. Thus, the summation over k products results in replacing the infinity distance by the smallest exponent summation until either the result is one (1) or k is exhausted.

5 At block 550, the routing calculation block 345 determines if the next_step cost matrix is equal to the matrix of the previous step. Accordingly, in the illustrated embodiment, at the block 550, the D_{ij}^2 cost matrix is compared to the previous cost matrix, D_{ij} . The two matrixes are equal if all of the corresponding elements of the matrices are equal. Since the D_{ij}^2 cost matrix and the D_{ij} matrix in the illustrated embodiment are not equal, at the block 530, the next_step cost matrix is re-calculated. If, at the block 550, the resultant matrix is equal to the matrix of the previous step, then it is an indication that no more processing is required because the resultant matrix represents the minimum distance between any two nodes for any given number of steps.

15 At the block 530, the next_step cost matrix (e.g., D_{ij}^3 cost matrix) is calculated using the following mathematical function:

$$D_{ij}^3 = \min_{k=1 \text{ to } 6} (D_{ik}, D_{kj}^2),$$

20 where i indicates the row and j indicates the column of the cost matrix (D_{ij}), D_{ij}^2 is the previous step matrix calculated above, and k ranges from 1 to the number of nodes in the network. The resultant next_step cost matrix in the illustrated embodiment is:

25

$$D_{ij}^3 = \begin{vmatrix} 1 & D^2 & D^5 & D^2 & \infty & \infty \\ \infty & 1 & D^3 & D^2 & \infty & \infty \\ \infty & \infty & 1 & \infty & D & D^5 \\ \infty & D^2 & D^3 & 1 & D & \infty \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} \begin{vmatrix} 1 & D^2 & D^4 & D^2 & D^2 & D^{10} \\ \infty & 1 & D^3 & D^2 & D^3 & D^8 \\ \infty & \infty & 1 & \infty & D & D^3 \\ \infty & D^2 & D^3 & 1 & D & D^3 \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} = \begin{vmatrix} 1 & D^2 & D^4 & D^2 & D^2 & D^4 \\ \infty & 1 & D^3 & D^2 & D^3 & D^5 \\ \infty & \infty & 1 & \infty & D & D^3 \\ \infty & D^2 & D^3 & 1 & D & D^3 \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix}$$

5 The resultant matrix, D_{ij}^3 , illustrates the minimum distance between any two nodes (node i (row) and node j (column)) in three or fewer steps. At the block 530, the resultant next_step cost matrix (e.g., D_{ij}^3 cost matrix) is compared to the previous step cost matrix (e.g., D_{ij}^2 cost matrix) to see if the two matrices are equal. And, since in the instant case the two matrices are not equal, the next_step cost matrix, at the block 530, is computed once again.

10 At the block 530, the next_step cost matrix (e.g., D_{ij}^4 cost matrix) is calculated using the following mathematical function:

15 where i indicates the row and j indicates the column of the cost matrix (D_{ij}), D_{ij}^3 is the previous step matrix calculated above, and k ranges from 1 to the number of nodes in the network. The resultant next_step cost matrix in the illustrated embodiment is:

$$D_{ij}^4 = \min_{k=1 \text{ to } 6} (D_{ij}, D_{kj}^3),$$

$$D_{ij}^4 = \begin{vmatrix} 1 & D^2 & D^5 & D & \infty & \infty \\ \infty & 1 & D^3 & D^2 & \infty & \infty \\ \infty & \infty & 1 & \infty & D & D^5 \\ \infty & D^2 & D^3 & 1 & D & \infty \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} \begin{vmatrix} 1 & D^2 & D^4 & D & D^2 & D^{10} \\ \infty & 1 & D^3 & D^2 & D^3 & D^8 \\ \infty & \infty & 1 & \infty & D & D^3 \\ \infty & D^2 & D^3 & 1 & D & D^3 \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix} = \begin{vmatrix} 1 & D^2 & D^4 & D & D^2 & D^4 \\ \infty & 1 & D^3 & D^2 & D^3 & D^5 \\ \infty & \infty & 1 & \infty & D & D^3 \\ \infty & D^2 & D^3 & 1 & D & D^3 \\ \infty & \infty & \infty & \infty & 1 & D^2 \\ \infty & \infty & \infty & \infty & \infty & 1 \end{vmatrix}$$

20 The resultant matrix, D_{ij}^4 , illustrates the minimum distance between any two nodes (node i (row) and node j (column)) in four or fewer steps. At the block 530, the resultant next_step cost matrix (e.g., D_{ij}^4 cost matrix) is compared to the previous step cost matrix (e.g., D_{ij}^3 cost matrix) to see if the two matrices are equal. And, since all of the elements of the two matrices are equal, it is an indication that previous step matrix, D_{ij}^3 , illustrates the minimum distance between any two nodes for any given number of steps.

Upon calculating the minimum distance between any two nodes within the network, the routing calculation block 345, at block 550, determines the efficient route or path the data packet 250 takes by determining any or all of the intermediate nodes between the source node (*e.g.*, first node 6) and destination node (*e.g.*, sixth node 11). One embodiment of the block 550 is illustrated in Figure 6. For illustrative purposes, the method of Figure 6 is described herein with reference to a specific example of determining the shortest distance from the first node 6 to the sixth node 11. The method of Figure 6 begins at block 610, where the routing calculation block 345 determines the minimum distance between a current pair of nodes. In one embodiment, during the first iteration, the current pair of nodes represents the source and destination nodes, which in the instant example includes the first and sixth nodes 6, 11. In one embodiment, the minimum distance is determined with reference to the element at row one, column six of the cost matrix, D_{ij}^3 , which, as stated earlier, illustrates the minimum distance between any two nodes for any given number of steps. In this example, the minimum distance is four (4) between the first and sixth nodes 6, 11.

At block 620, the routing calculation block 345 determines at least one route/path having the minimum distance of four between the current pair (*e.g.*, first and sixth nodes 6,11) of nodes. In one embodiment, to determine the route/path having the minimum distance of four, the IP route block 550 employs the mathematical function shown in equation (1) below.

25

$$D^3(F, S)_{\min} = \min_{k=1 \text{ to } 6} (D^3(F)(k) * D^3(k)(S)), \quad (1)$$

where F and S represent respective first and second input nodes, and D_{ij}^3 is the cost matrix illustrating the minimum distance between any two nodes for any given number of steps. Equation (1) essentially determines the minimum

distance between a source node and destination node and all of the intermediate nodes. Thus, in the illustrative example (where $F=1$ and $S=6$), equation (1) essentially determines the distance from the first node 6 to the sixth node 11, then distance from the first node 6 to the second node 7 and between the 5 second node 7 to the sixth node 11, then the distance from the first node 6 to the third node 8 and third node 8 to the sixth node 11, then the distance from the first node 6 to the fourth node 9 and fourth node 9 to the sixth node 11, and so forth.

10 As mentioned above, in the illustrative example, F is one, and S is six. Thus, employing the mathematical expression of equation (1) to the illustrated example yields the terms $[(1)(D^4)=D^4, (D^2)(D^5)=D^7, (D^4)(D^3)=D^7, (D)(D^3)=D^4, (D^2)(D^2)=D^4, (D^4)(D^1)=D^4]$, which indicates that, aside from the first and sixth nodes 6, 11, the fourth and fifth nodes 9, 10 involve the shortest distance, 15 namely 4. This means that that two possible minimum-distance routes from the first node 6 to the sixth node 11 are either: $1 \rightarrow 5 \rightarrow 6$ OR $1 \rightarrow 4 \rightarrow 6$.

20 Assuming that the route/path under consideration is $1 \rightarrow 5 \rightarrow 6$, at block 640, the routing calculation block 345 determines if a direct connection exists between a first pair of nodes in the at least one route (*e.g.*, $1 \rightarrow 5 \rightarrow 6$). In the instant example, the first pair of nodes includes the first and fifth nodes 6, 10. In one embodiment, the connection matrix (C) depicted above is utilized to 25 determine if a direct connection exists between the first pair. The element at row one, column five of the connection matrix is zero, which indicates that no direct connection exists between the first and fifth nodes 6, 10. Thus, at block 650, the routing calculation block 345 uses the next route (*e.g.*, $1 \rightarrow 4 \rightarrow 6$) of the at least one route that was calculated at the block 620.

30 Assuming that the route/path now under consideration is $1 \rightarrow 4 \rightarrow 6$, at the block 640, the routing calculation block 345 determines if a direct connection

exists between a first pair of nodes in the at least one route (e.g., 1 → 4 → 6). In the instant example, the first pair of nodes includes the first and fourth nodes 6, 9. Using the connection matrix, the routing calculation block 345 determines that a direct connection exists between the first and fourth nodes 6, 9 since the 5 element at row one, column four is one.

At block 660, the routing calculation block 345 determines if any additional pair of nodes exists in the at least one route (e.g., 1 → 4 → 6) calculated at the block 620. Since another pair (4 and 6) of nodes does exist in 10 the at least one route, at block 670, the routing calculation block 345 determines if a direct connection exists between the fourth and sixth nodes 9, 11. The routing calculation block 345 checks to see the value of the element at row four, column six to determine if a direct connection exists between the fourth and sixth nodes 9, 11. Since the value of element at the row four, column six is zero, 15 no direct connection exists between these two nodes, which means that, at block 680, the variable "current pair of nodes" is set to equal the next pair of nodes (e.g., 4 and 6), and the process of Figure 6 is once again repeated.

At the block 610, the routing calculation block 345 determines the minimum distance between the current pair of nodes (e.g., nodes 4 and 6). The minimum distance is determined with reference to the element at row four, column six of the cost matrix, D^3_{ij} , which, as mentioned before, illustrates the minimum distance between any two nodes for any given number of steps. In this example, the minimum distance is four (3) between the fourth and sixth 25 node 9, 11.

At the block 620, the routing calculation block 345 determines at least one route/path having the minimum distance of three between the current pair of nodes (e.g., fourth and sixth nodes 9,11). In one embodiment, to determine the route/path having the minimum distance of three, the routing calculation block 30

345 employs the mathematical function shown above in equation (1). In this case, F is four, and S is six. Thus, employing the mathematical expression of equation (1) to the illustrated example yields the terms $[(\infty)(D^4)=\infty, (D^2)(D^5)=D^7, (D^3)(D^3)=D^6, (1)(D^3)=D^3, (D)(D^2)=D^3, (D^3)(1)=D^3]$, which indicates that, aside from the fourth and sixth nodes 9, 11, the fifth node 10 involves the shortest distance, namely 3. This means that the minimum-distance route from the fourth node 9 to the sixth node 11 is $4 \rightarrow 5 \rightarrow 6$.

Assuming that the route/path now under consideration is $4 \rightarrow 5 \rightarrow 6$, at the block 640, the routing calculation block 345 determines if a direct connection exists between a first pair of nodes in the at least one route (*e.g.*, $4 \rightarrow 5 \rightarrow 6$). In the instant example, the first pair of nodes includes the fourth and fifth nodes 9, 10. Using the connection matrix, the routing calculation block 345 determines that a direct connection exists between the fourth and fifth nodes 9, 10, since the element at row four, column five is one.

At the block 660, the routing calculation block 345 determines if any additional pair of nodes exists in the at least one route (*e.g.*, $4 \rightarrow 5 \rightarrow 6$) that was calculated at the block 620. Since another pair (5 and 6) of nodes exist in the at least one route, at the block 670, the routing calculation block 345 determines if a direct connection exists between the fifth and sixth nodes 10, 11. Since the element at row five, column six of the connection matrix is one, the route calculation block 345 determines that a direct connection exists between the fifth and sixth nodes 10, 11.

At the block 660, the routing calculation block 345 again determines if any more pair of nodes exists in the at least one route (*e.g.*, $4 \rightarrow 5 \rightarrow 6$). In the instant case since no more pairs exits, at block 690, the routing calculation block 345 calculates, based on the earlier iterations, the minimum distance route/path between the first and sixth nodes 6, 11. That is, in the instant example, it is

determined that the optimal route in terms of the minimum distance (*e.g.*, 4) between the first node 6 and sixth node 11 is from the first node 6 to the fourth node 9 to the fifth node 10 to the sixth node 11 (*e.g.*, 1 → 4 → 5 → 6).

5 Referring again to Figure 4, at block 480, the IP output block 300 updates the IP options field 274 of the data packet 250 based on the efficient route calculated at the block 470. Thus, in the illustrated embodiment, since the shortest distance from the first host 4 to the second host 5 is through the first node, fourth node, fifth node, and sixth node, the IP output block 300 of the first router updates the IP options field 274 with the IP addresses corresponding to the fourth, fifth, and sixth nodes 9, 10, 11. In one embodiment, upon writing to the IP options field 274, the IP output block 300 of the first router 26 indicates that source routing is desired.

15 At block 490, the data packet 250 is transmitted to the next router, which in the illustrated embodiment is the fourth router 29. When the data packet 250 reaches the fourth router 29, it is queued and then processed by the IP options processing block 295 of the fourth router 29. And since the first router 26, at the block 480, indicates that source routing is desired, the IP options processing block 295 of the fourth router 29 provides the data packet 250 to its IP output block 300. The IP output block 300 then updates the IP options field 274 to reflect that the next destination of the data packet 250 is the fifth node 10. The above-described process continues until the data packet 250 is ultimately delivered to the second host 5.

20
25 In accordance with one embodiment of the present invention, instead of using source routing, each router that receives the data packet 250 may independently calculate a cost efficient path using the method of Figure 5 for delivering the data packet 250 to its final destination. That is, upon calculating the next node, the router transmits the data packet 250 without indicating the

use of source routing, thereby leaving it up to the receiving router to calculate the next destination. Without the use of source routing, each router may independently use the method of Figure 5 for calculating an efficient route to deliver the data packet 250 to its final destination.

5

In accordance with one embodiment of the present invention, instead of determining the connection and/or cost matrices, the router 200 (see Figure 2) may simply receive, and perhaps store, the matrices that might have been determined elsewhere. Thus, in one embodiment, upon receiving the connection 10 and/or cost matrices, the router 200 calculates a cost efficient route from a source node to a destination node.

15 Although in the illustrated embodiment the connection and cost matrices are generated based on adjacent nodes, such matrices are not limited as such. That is, the connection and/or cost matrices may be based on any combination of nodes or routers (*e.g.*, even/odd nodes or routers) in the network 3.

20 Some embodiments may have one or more of the following advantages. A reliable and efficient method and apparatus are provided for routing signals from a source to a destination node. In one embodiment, a data packet may be transmitted to its intended destination in a "cost" efficient manner, where the cost may be one or more of the factors such as the physical distance, delay, expense, reliability, bandwidth, load, and security between nodes. Accordingly, 25 with the advent of the present invention, routers may, in one or more embodiments, transmit data packets to their final destination expeditiously and accurately, thereby improving network congestion.

30 The various system layers, routines, or modules may be executable control units (such as control unit 355) (see Figure 2). Each control unit may include a microprocessor, a microcontroller, a processor card (including one or

more microprocessors or controllers), or other control or computing devices. The storage devices referred to in this discussion may include one or more machine-readable storage media for storing data and instructions. The storage media may include different forms of memory including semiconductor memory devices such as dynamic or static random access memories (DRAMs or SRAMs), erasable and programmable read-only memories (EPROMs), electrically erasable and programmable read-only memories (EEPROMs) and flash memories; magnetic disks such as fixed, floppy, removable disks; other magnetic media including tape; and optical media such as compact disks (CDs) or digital video disks (DVDs). Instructions that make up the various software layers, routines, or modules in the various systems may be stored in respective storage devices. The instructions when executed by a respective control unit cause the corresponding system to perform programmed acts.

In the loading or transport process, data signals that are embodied in carrier waves (transmitted over telephone lines, network lines, wireless links, cables, and the like) may communicate the code segments, including instructions, to the terminal. Such carrier waves may be in the form of electrical, optical, acoustical, electromagnetic, or other types of signals.

While the invention has been disclosed with respect to a limited number of embodiments, those skilled in the art will appreciate numerous modifications and variations therefrom. It is intended that the appended claims cover such modifications and variations as fall within the true spirit and scope of the invention.